NASHVILLE SOFTWARE SCHOOL

Project 1

Data engineering

Introduction

This first project will have three parts. You will first get data from an api and store it, then you will get data from a csv and store it, and you will create a presentation for the class.

(Stretch goal) lastly, if time permits, you will present a case for how you would handle data integrity. (making sure the data you read, is the same data you wrote)

Part 1

- You will be given a Business requirements document (BRD) with the following
    - Business use case
    - Documentation for the API
- You will need to create an ER model, and then implement the ER model in a postgres database.
- You will need to access the data from the API
    - This data will be in JSON format
- You will need to transform the Json data to match your table schema
- You will need to load the data into your postgres database

Part 2

- You will be given a file in csv format
- You will need to create an ER model, and then implement the ER model in a postgres database
- You will process the file and load the data into your postgres database
    - The data may have null values

Part 3 (Stretch Goal)

- Data Validation
    - You will need to come up with a way to validate that the data that you wrote to the database matches the data was in the csv file.
    - This is important, as you may be dealing with financial transactions in the future and need to make sure the data that you move is accurate.

Presentation

- You will create a slide deck
    - Talk about Architectural Decisions
    - Talk about any challenges you had to overcome.
    - Talk about anything that you would do differently
    - Address anything that you tried, but could not get to work, if anything.

Summary

Part 1: Data from API to database

Part 2: Data from csv to database

Part 3: Presentation

Stretch Goal: Data validation.